



www.turkishstudies.net/appliedsciences

Turkish Studies - Information Technologies and Applied Sciences

eISSN: 2667-5633

Research Article / Araştırma Makalesi



INTERNATIONAL
BALKAN
UNIVERSITY
Sponsored by IBU

Yapay Zekâ ile Konut Fiyatlarının Tahmin Edilmesi

Estimation of Housing Prices with Artificial Intelligence

Emrah Aydemir* - Cemal Aktürk** - Mehmet Ali Yalçinkaya***

Abstract: Housing purchase is a multidimensional problem that requires people to meet their housing, health, safety and various socio-cultural needs with a certain budget. The developments in artificial intelligence methods now allow solutions to many problems such as buying houses in daily life. Artificial neural networks and deep learning methods are among the methods used for this purpose. In this study, an intelligent system has been developed that collects and records the housing data for sale on the internet and estimates the housing price. Housing data for sale in Istanbul was collected for the developed system. The data were divided into 10-fold cross-validation method for training and testing. The data were modeled with 14 different algorithms based on 176 attributes, and price estimation study was performed. The most successful price estimation for 852 houses in Ataşehir district of Istanbul was obtained by Random Forest Algorithm. With this study, a system has been developed in which researchers, companies operating in the housing market and consumers can estimate the housing prices. In addition, it is aimed to give researchers a new perspective for using different artificial intelligence methods or residential attributes.

Structured Abstract: Throughout history, the main task of houses has been to protect people from physical factors. Nowadays, besides meeting the housing needs of people, it also provides the opportunity to invest and earn money. Houses have also become the most clearly reflecting places of people's social and cultural classes. The mentioned factors have resulted in a serious growth of the housing trading market. In addition, factors such as the increase in the number of family members and the insufficiency of existing houses to the needs caused people to search for new houses. Today, the price of the houses is of great importance as well as the properties of the houses in the process of housing purchase. Housing prices vary due to the free market

* Dr. Öğr. Üyesi, Kırşehir Ahi Evran Üniversitesi, Mühendislik-Mimarlık Fakültesi, Bilgisayar Mühendisliği Bölümü
Asst. Prof. Dr., Kırşehir Ahi Evran University, Faculty of Engineering-Architecture, Department of Computer Engineering
ORCID 0000-0002-8380-7891

emrah.aydemir@ahievran.edu.tr

** Dr. Öğr. Üyesi, Gaziantep İslam Bilim ve Teknoloji Üniversitesi, Mühendislik ve Doğa Bilimleri Fakültesi, Bilgisayar Mühendisliği Bölümü

Asst. Prof. Dr., Gaziantep İslam Science and Technology University, Faculty of Engineering and Natural Sciences, Department of Computer Engineering

ORCID 0000-0003-3764-3862

cemalakturk79@gmail.com

*** Arş. Gör. Dr. Kırşehir Ahi Evran Üniversitesi, Mühendislik-Mimarlık Fakültesi, Bilgisayar Mühendisliği Bölümü
Res. Asst. Dr., Kırşehir Ahi Evran University, Faculty of Engineering-Architecture, Department of Computer Engineering
ORCID 0000-0002-7320-5643

mehmetyalcinkaya@ahievran.edu.tr

Cite as/ Atıf: Aydemir, E., Aktürk, C., Yalçinkaya, M. A. (2020). Yapay zekâ ile konut fiyatlarının tahmin edilmesi. *Turkish Studies - Applied Sciences*, 15(2), 183-194. <https://dx.doi.org/10.29228/TurkishStudies.43161>

Received/Geliş: 25 April/Nisan 2020

Accepted/Kabul: 20 June/Haziran 2020

Copyright © INTAC LTD, Turkey

Checked by plagiarism software

Published/Yayın: 25 June/Haziran 2020

CC BY-NC 4.0

economy, and large price differences can occur even among similar houses in the same city. For this reason, it is of great importance to set realistic prices for houses. Because people spend their years of savings, or most of their future earnings in their homes. This is a one-off procedure for many families today, with no luxury to make mistakes. In making such an important decision, the systems that make an accurate price estimation according to the house to be taken are of great importance.

The random forest algorithm is a batch classification algorithm developed by Leo Breiman and Adele Cutler, which includes the voting method. A decision forest developed with a random forest algorithm is formed by combining many decision trees. Class estimation process with random forest algorithm; It is determined by voting by the trees in the forest. Decision trees in the developed forest are created using examples separated from the dataset using the bootstrap technique. Classification and forecasting processes using random forest algorithm basically consist of 4 steps. In the first step, n sampling is selected from the original data set. $2/3$ of each selected sample is used as training and $1/3$ as a testing set. In the second step, among the variables randomly selected from the training set, the variable with the best information gain is selected as a branch separator. This process is also known as Classification and Regression Tree (CART). Random Forest algorithm is preferred more than collective learning algorithms because it provides high accuracy rate. The random forest algorithm is more advantageous than the Boosting and Bagging methods thanks to its fast operation and high accuracy rate. Deep learning; obtains the complexity of multidimensional data sets by using back propagation algorithm. In deep learning, each layer takes the output of the previous layer as input. To put it in more detail; each layer first calculates the parameter values and then compares with the values obtained from the previous layer and decides the necessary changes. The basis of deep learning is based on multi-layer artificial neural networks (ANN).

In this study, an intelligent system has been developed that collects and records the housing data for sale on the internet and estimates the housing price. In this study, a system has been developed that determines the estimated price by using the data of apartment advertisements for sale in Ataşehir district of Istanbul. In this section, artificial intelligence methods used in the development of the system in question are examined. 852 housing data belonging to Ataşehir district of Istanbul province were collected from the internet within the scope of the study. Housing prices vary between 142.500 TL and 1.000.000 TL. Houses vary between 37 m² and 300 m² gross. While housing prices for 16 districts in total are concentrated between 300,000 TL and 600,000 TL, houses with more than a million liras are excluded. The majority of the residences consist of 2 + 1 and 3 + 1 flats and the building ages are mostly zero or between 5-10 years old. Apartments for sale are mostly apartments in the range of 1-5, and garden floor and below. While the clear majority of the apartments are heated by the natural gas combi boiler, most the remaining ones have a heat share meter system. Housing prices in Ataşehir district were estimated by various algorithms using the Weka program. The data were modeled with 14 different algorithms based on 176 attributes, and price estimation study was performed. The most successful price estimation for 852 houses in Ataşehir district of Istanbul was obtained by Random Forest Algorithm. Because of the analysis, the most successful result was obtained with the random forest algorithm. The data were divided into 10-fold cross-validation method for training and testing. Housing prices can be estimated with an average difference of ± 80 thousand TL. The correlation coefficient between real prices and estimated prices was found to be 0.86. This shows that there is a high relationship between the two and shows that a successful prediction has been achieved. With this study, a system has been developed in which researchers, firms operating in the housing market and consumers can estimate the housing prices. In addition, it is aimed to give researchers a novel perspective for using different artificial intelligence methods or residential attributes.

Keywords: House Price Estimation, Decision Support System, Random Forest Algorithm

Öz: Konut satın alma, insanların belirli bir bütçe ile barınma, sağlık, güvenlik ve çeşitli sosyokültürel ihtiyaçlarını karşılamayı gerektiren çok boyutlu bir problemdir. Yapay zekâ yöntemlerinde yaşanan gelişmeler, artık günlük hayatta konut satın alma gibi birçok problemin çözümüne imkân tanımaktadır. Yapay sinir ağları ve derin öğrenme yöntemleri bu amaçla kullanılan yöntemlerdendir. Bu çalışmada, internetteki satılık konut verilerini toplayarak kaydeden ve konut fiyat tahmini yapan zeki bir sistem geliştirilmiştir. Geliştirilen sistemde kullanılmak üzere İstanbul iline ait satılık konut verileri toplanmıştır. Veriler 10-katlı çapraz doğrulama yöntemi ile eğitim ve test olacak şekilde ayrıştırılmıştır. Veriler 176 özneliğe göre 14 farklı algoritma ile modellenerek fiyat tahmin çalışması yapılmıştır. İstanbul'un Ataşehir ilçesindeki 852 konut için en başarılı

fiyat tahmini Rastgele Orman Algoritması ile elde edilmiştir. Yapılan çalışma ile araştırmacıların, konut piyasasında faaliyet gösteren firmaların ve tüketicilerin konut fiyatlarını tahmin edebileceği bir sistem geliştirilmiştir. Ayrıca, sonraki çalışmalarda farklı yapay zekâ yöntemlerinin veya konut özniteliklerinin kullanılması için araştırmacılara yeni bir bakış açısı kazandırılması hedeflenmiştir.

Anahtar Kelimeler: Konut Fiyat Tahmini, Karar Destek Sistemi, Rastgele Orman Algoritması

Giriş

İnsanların yaşamını sürdürmesi için beslenme, barınma ve korunma gibi bir takım temel ihtiyaçlarını karşılaması gerekmektedir. Hayatın olağan akışı içerisinde bireyler, eğitimlerini ve kişisel gelişimlerini tamamladıktan sonra kendi ailelerini kurabilmek için bir işe ve sonrasında aileleriyle birlikte yaşayacakları bir eve ihtiyaç duyarlar. Mesleki gelişimini ve dolayısıyla ekonomik gelişimini belirli bir olgunluk seviyesine ulaştırmayı başaran insanlar, doğal bir hedef olarak artık aileleriyle kendi evlerinde oturmanın planını yaparak bir konut satın alma motivasyonu ile hareket ederler. Ebeveynlerin ve diğer aile fertlerinin tüm sosyal ihtiyaçlarını karşılamak amacıyla mevcut bütçeyi de göz önünde bulundurarak satın alınacak konut konusunda doğru kararı vermek, konut alma probleminin bağlı olduğu değişkenler açısından karmaşık bir süreçtir. Çünkü konuta ayrılan bütçe ile alınacak konutun beklentileri karşılaması her zaman mümkün değildir.

Konutlar her ne kadar insanların barınma sorununa karşın ortaya çıkmış bir ihtiyaç ve tüketim malı gibi görünse de aynı zamanda yerli paranın değer kaybına karşı korunmak amacıyla bir yatırım aracı olarak da görülmektedir. Bu sebeple konutlar sadece sosyal bir ihtiyaç olmayıp, aynı zamanda ekonomik, kültürel, hukuksal ve teknolojik boyutları da olan bir üründür. Bu boyutları açıklamak gerekirse, insanların yaşam tarzı ve kültürel ihtiyaçlarına karşı konut ve konutun bulunduğu mekân ile ilgili tercihleri de değişmektedir. Aynı zamanda bir konuta sahip olurken bunun hukuki olarak da güvence altına alınması gerekmektedir. Konutlar; gelişen teknoloji ve ortaya çıkan Endüstri 4.0 trendindeki nesnelerin interneti kavramıyla birlikte akıllı evlere dönüşmekte, bu sayede güncel ihtiyaçlara da cevap vermektedir. Tüm bu hususlara ilaveten konut piyasası; insanların gelir seviyesi, medeni durumu, toplumun sanayileşmesi ve tarımsal istihdam oranı, faiz oranları, nüfus artışı ve göç gibi birçok değişkenden etkilenir. Fitöz ve Öztürk (2008) yaptıkları çalışmada; kişi başına düşen milli gelir, konut fiyatları ve faiz oranları ile konut talebi arasında olumlu bir ilişki olduğunu göstermiştir. Lebe ve Akbaş (2014) da çalışmalarında aynı değişkenlerle konut talebi arasında uzun dönemli bir ilişkinin bulunduğunu tespit etmiştir. Arslan ve arkadaşları (2013), 2025 yılına kadar sadece nüfus artışı ve nüfus yapısındaki değişikliklerden dolayı konut talebinin %1,65 artacağını belirtmiştir. Bu sebeplerle bir konutun fiyatı belirlenirken bir ürün olarak konutun girdi maliyetlerine ek olarak bahsedilen tüm değişkenler de fiyata etki etmektedir.

Konutun bulunduğu mevki de fiyata ve konut tercihine etki eden bir diğer önemli değişkendir (Türel, 1976). Özellikle çalışılan iş yeri ile konutun arasındaki ulaşım imkânları ve bu ulaşımın geçen süre, konut yeri seçiminde etkili olmaktadır (Marin ve Altıntaş, 2014). Ayrıca konutun bulunduğu mekânın estetiği bile konutu satın alma veya kiralama talebini ve konutun değerini etkilemektedir (Ekşioğlu ve Çubukçu, 2011). Bu sebeple bir konutun alan genişliği, oda sayısı, banyo sayısı gibi yapısal özelliklerine ek olarak bulunduğu il, ilçe, semt gibi konum özellikleri de konutun fiyatı ve sosyal ihtiyaçları tatmin etme açısından önem arz etmektedir. Örneğin İstanbul ilindeki bir konut yerinin ortalama metrekare fiyatı ile daha küçük bir il olarak Bartın ilindeki bir arsanın ortalama metrekare fiyatları aynı olmayacaktır. Arsa fiyatlarında illere göre değişiklik görüldüğü gibi aynı ilin ilçelerine göre de farklılıklar görülmesi tabiidir. İstanbul'da yerleşime ve gelişime yeni başlayan Arnavutköy ilçesi ile Beşiktaş, Sarıyer gibi daha merkezi ilçelerin arsa fiyatlarının farklı olması da bu ilçelerdeki konut talebini ve dolayısıyla konut fiyatlarını farklılaştıracaktır. Konutun bulunduğu konumun arsa değeri aynı zamanda şehir merkezine ulaşım imkânlarına, okul, hastane ve alışveriş merkezi gibi sosyal ihtiyaçların giderileceği mekânlara yakınlık durumuna göre de

değişkenlik gösterecektir. Toplu taşıma ulaşım hatlarına ve duraklarına daha yakın konutlar, emsallerine göre daha değerli olacağından daha yüksek fiyatlı olacaktır.

Bu çalışmada, internetteki konut ilanlarından otomatik veri toplayan bir sistem geliştirilmiştir. Geliştirilen sistemde konutların oda sayısı, alan genişliği, bulunduğu kat, banyo sayısı, bina yaşı vb. yapısal özellikleri ile eğitim, sağlık ve sosyal ihtiyaç yerlerine, şehir merkezi ve toplu ulaşım hatlarına mekânsal yakınlıklarını belirten özelliklerden oluşan toplam 176 öznelik belirlenmiştir. Geliştirilen sistem ile İstanbul'un Ataşehir ilçesine ait 2019 yılının Kasım ve Aralık aylarında yayınlanan 852 adet satılık konut ilan verileri kullanılmıştır. Toplam 14 farklı algoritma ile konut fiyatları için tahmin çalışması gerçekleştirmiştir. Yapılan çalışma ile geliştirilen sistemin, hem satılık konut ilanlarını otomatik toplanması hem de 176 konut özneliği için 14 farklı algoritma ile fiyat tahmin çalışmasını gerçekleştirerek Türkiye'nin en büyük kentine örnek olarak uygulanması amaçlanmıştır. Bu açıdan konut fiyat tahmininde veri toplama, kullanılan öznelik sayısı ve yapay zekâ teknikleri açısından yapılan çalışma ile literatüre yeni bir bakış açısı kazandırılması amaçlanmıştır.

Konut fiyat tahminine yönelik yapılan çalışmalar incelendiğinde, çalışmaların genellikle konut maliyetlerinin tahminine yönelik, regresyon ve lineer yöntemlerin kullanıldığı, hedonik yöntem ile şehirlerdeki konut fiyatlarının tahmin edildiği ve yapay sinir ağlarının kullanıldığı tahmin çalışmaları olduğu görülmektedir. Erdem (1996), binaların fonksiyonel alt elemanlarına ayrıştırılarak inşaatın birim metrekare maliyetini tahmin etmek için bir maliyet tahmin modeli geliştirmiştir. Bostancıoğlu (1999, 2006) yaptığı çalışmalarda konut maliyetine etki eden faktörleri tanımlayarak bu faktörlerin yeni bir proje için uyarlandığında oluşacak konut maliyetini tahmin etmiştir. Uğur ve Sivri (2014) toplu konut projelerindeki inşaat maliyetlerini regresyon yöntemi ile inceleyerek çeşitli yıllara göre konut maliyetlerindeki değişimi göstermiştir. Tuna ve arkadaşları (2015) Ankara ilindeki satılık konut ilanlarını inceleyerek çoklu regresyon analizi ve coğrafi bilgi sistemleri yardımıyla ev fiyatları için bir tahmin çalışması yapmıştır. Araştırmacılara göre Ankara'daki evlerde ankastre mutfak veya merkezi ısıtma sisteminin olması ev fiyatını en çok arttıran değişkenlerdir. Ayrıca araştırmacılar bina yaşının artmasının evin fiyatını düşürmediği, aksine daha merkezi konumda olan ancak eski olan evlerin konumundan dolayı daha yüksek fiyatlarda olduğunu vurgulamıştır.

Uygar ve Yayla (2016) İstanbul, Hayrullahoğlu ve arkadaşları (2017) ile Gökler (2017) Ankara, Ekşioğlu ve Çubukçu (2011) ile Ecer (2014) İzmir, Kördiş ve arkadaşları (2014) Antalya, Yayar ve Gül (2014) Mersin, Afşar ve arkadaşları (2017) Eskişehir, Yayar ve Karaca (2014) Tokat, Samsun, Çorum ve Amasya illeri için mekânsal bağımlılığı dikkate alan hedonik konut fiyatlama modelini kullanarak konut fiyatlarını tahmin etmişlerdir. Ecer (2014) ise çalışmasında konut fiyat tahmini için hedonik model ile yapay sinir ağlarını karşılaştırarak yapay sinir ağlarının konut fiyatını tahmin etmede hedonik modelden daha başarılı olduğunu vurgulamıştır.

Yapay sinir ağları ile yapılan ilgili diğer çalışmalara bakıldığında; Demir (2007) toplu konutlardaki inşaat maliyetlerinin tahminini hesaplamıştır. Uğur ve arkadaşları (2011) tek katlı yığma konut maliyetlerini yapay sinir ağlarını kullanarak tahmin etmiştir. Yılmazel ve arkadaşları (2018) ise Eskişehir ilindeki satılık konutlardan oluşan veri kümesinde yapay sinir ağlarını kullanarak konutların özellikleri ile konut fiyatları arasındaki ilişkiyi incelemiştir.

Yöntem

Veri Seti

Konut gayrimenkulü birçok farklı özelliğe sahip olan bir mal olduğundan dolayı, fiyatının belirlenmesinde söz konusu özellikler büyük önem taşımaktadır. Konut alım satım işlemlerinde fiyat belirleyici kesin kuralların olmaması, aynı bölgede yer alan konutların alım satım işlemlerinde dahi ciddi fiyat farklarının ortaya çıkmasına neden olmaktadır. Bu durum, konutun sahip olduğu özelliklere göre, fiyat tahmini yapacak sistemlere olan ihtiyacı arttırmaktadır. Bu çalışmada İstanbul

ilinin Ataşehir ilçesindeki satılık daire ilanlarına ait veriler kullanılarak tahmini fiyat belirlemesi yapan bir sistem geliştirilmiştir. Bu bölümde söz konusu sistemin geliştirilmesinde kullanılan yapay zekâ metotları incelenmektedir.

Rastgele Orman Algoritması

Toplu sınıflandırma algoritmaları; tek sınıflandırıcı yerine çok sayıda sınıflandırıcı kullanan ve bu sınıflandırıcılardan elde edilen tahmin sonuçlarını birleştirerek veriyi sınıflandıran algoritmalar. Günümüzde yaygın olarak kullanılmakta olan toplu sınıflandırma algoritmaları; Boosting (Hızlandırma), Bagging (Toralama) ve Random Forest (Rastgele Orman) algoritmalarıdır. Rastgele orman algoritmasına değinmeden önce, diğer toplu sınıflandırma algoritmalarına değinilmesi gerekmektedir.

Toplu sınıflandırma algoritmalarından bagging (torbalama) algoritmasında, eğitim setinden yer değıştırme işlemi ile çok sayıda önyüklemeli eğitim setleri oluşturulmaktadır. Oluşturulan her bir eğitim seti için bir ağaç üretilmektedir. Üretilen ağaçlardan elde edilen tahmin değeri içerisinde en çok oyu alan tahmin değeri sınıflandırma sonucu olarak belirlenmektedir. Diğer bir toplu sınıflandırma algoritması olan boosting (hızlandırma) algoritmasında iteratif olarak tekrarlı eğitim işlemi gerçekleştirilmekte, yanlış sınıflandırılan örneklerin ağırlıkları artırılarak sonraki iterasyonda tekrar sınıflandırılmaktadır. Hızlandırma algoritması, torbalama algoritmasına göre daha doğru sonuçlar vermesine rağmen, daha yavaş çalışmaktadır (Akar ve Güngör, 2012).

Rastgele orman algoritması, Leo Breiman ve Adele Cutler tarafından geliştirilen, içerisinde oylama metodunu barındıran bir toplu sınıflandırma algoritmasıdır. Rastgele orman algoritması ile geliştirilmiş bir karar ormanı, birçok karar ağacının bir araya gelmesi ile oluşmaktadır. Rastgele orman algoritması ile sınıf tahmin işlemi; oluşturulan orman içerisindeki ağaçlar tarafından oylanarak belirlenmektedir. Geliştirilen orman içerisinde yer alan karar ağaçları, veri setinden bootstrap tekniği ile ayrılan örnekler kullanılarak oluşturulmaktadır. Rastgele orman algoritması kullanılarak gerçekleştirilen sınıflandırma ve tahmin işlemleri temelde 4 adımdan oluşmaktadır. İlk adımda, orijinal veri setinden n adet örnekleme seçilmektedir. Seçilen her bir örnekleme $2/3$ 'ü eğitim (inBag), $1/3$ 'ü ise test (Out of Bag) seti olarak kullanılmaktadır. İkinci adımda, eğitim seti içerisinde rastgele seçilen değişkenler arasında en iyi bilgi kazancına sahip değişken, dallara ayırıcı olarak seçilmektedir. Bu işlem aynı zamanda Classification and Regression Tree (CART) olarak bilinmektedir. Dallara ayırma kriteri GINI indeksi ile belirlenmektedir. GINI indeksinin hesaplanmasında kullanılan formül, Denklem 1'de gösterilmektedir.

$$\sum_{j \neq i} \sum (f(C_i, T)/|T|)(f(C_j, T)/|T|) \quad (1)$$

Yukarıda gösterilen denklemde; $f(C_i, T)/|T|$ ifadesi seçilen örneğin C_i sınıfına ait olma olasılığını göstermektedir (Akar vd., 2010). Algoritmanın 3. Aşamasında gelindiğinde, yaprak düğüm elde edilene kadar bu işlem tekrar edilmektedir. Yaprak düğümün elde edilmesi için 2 temel şart bulunmaktadır. Bunlardan ilki, düğüm içerisindeki tüm örneklerin aynı sınıfa ait olması, ikincisi de bölünmenin yapılacağı değişken kalmamış olmasıdır. Algoritmanın 4. Ve son adımında; ilk adımda oluşturulan eğitim seti kullanılarak geliştirilen ağaca, yine aynı adımda oluşturulan test seti uygulanmaktadır. Bu işlem ile ağacın hata oranı hesaplanmaktadır. Oluşturulan modele sınıflandırılması için yeni bir veri gönderildiğinde, n örneklemeden oluşturulmuş n tane ağacın ürettiği tahminler birleştirilerek oylanmaktadır. Gerçekleştirilen oylama sonrasında en çok oyu alan sınıf, söz konusu veri için tahmin sınıfı olmaktadır (Korkem, 2013).

Rastgele orman algoritmasında, geliştirilen ormanın hata oranı 2 temel faktöre dayanmaktadır. Bunlardan ilki, orman içerisindeki ağaçlar arasındaki korelasyondur. Ağaçlar arasındaki korelasyonun artması, hata oranını arttırmaktadır. İkinci faktör ise, orman içerisinde yer alan ağaçların gücüdür. Ağaçların gücünün artırılması, hata oranının düşmesine neden olacaktır.

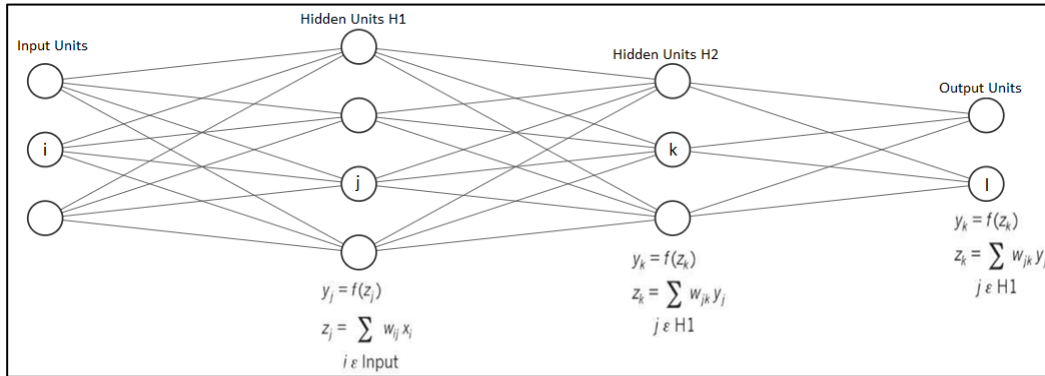
Ağaçlara ait korelasyon ve güç değerleri, girdi değişken sayısı (m)' e bağlı olarak değişmektedir. Girdi değişken sayısının azaltılması hem korelasyon hem de güç değerlerini azaltırken, artırılması söz konusu değerlerin artmasına neden olmaktadır (Ercire, 2019).

Rastgele Orman algoritması, yüksek doğruluk oranı sağladığı için toplu öğrenme algoritmalarına göre daha fazla tercih edilmektedir. Rastgele orman algoritması hem hızlı çalışması hem de yüksek doğruluk oranı sağlaması sayesinde Hızlandırma (Boosting) ve Torbalama (Bagging) yöntemlerine göre daha avantajlıdır (Akar ve Güngör, 2013).

Derin Öğrenme Algoritmaları

İlk olarak 1960'lı yıllarda ortaya çıkan derin öğrenme; yapay sinir ağları algoritmalarını kullanarak, çok boyutlu veriler üzerinde sınıflandırma ve tahmin işlemlerinin yapılmasına olanak sağlayan bir makine öğrenmesi alt dalıdır (Aktaş, 2020). Derin öğrenme algoritmaları sayesinde, görüntü işleme, doğal dil işleme, sınıflandırma vb. alanlarda elde edilen başarı oranları ciddi artış göstermiştir (Lecun vd., 2015).

Derin öğrenme; çok boyutlu veri setlerinin karmaşıklık yapısını geri yayılım algoritmasını kullanarak elde etmektedir. Derin öğrenmede her bir katman, kendinden önceki katmanın çıktısını girdi olarak almaktadır. Daha detaylı olarak ifade etmek gerekirse; her katman ilk olarak parametre değerlerini hesaplamakta, daha sonra da bir önceki katmandan elde edilen değerler ile karşılaştırma yaparak gerekli değişikliklere karar vermektedir. Derin öğrenmenin temeli, çok katmanlı yapay sinir ağlarına (YSA) dayanmaktadır. Şekil 1'de çok katmanlı bir yapay sinir ağı modeli gösterilmektedir. Şekilde gösterilmekte olan model incelendiğinde, nöronlar arası her bir bağlantı w_{ij} , w_{jk} , w_{kl} ağırlıkları ile ilişkilendirilmişlerdir (Aktaş, 2020).



Şekil 1: Yapay Sinir Ağı Yapısı

Derin öğrenme algoritmalarının en önemli özelliği, problemlere uygun özelliklerin belirlenmesinde sunduğu kolaylıktır. Derin öğrenme yöntemi ile oluşturulmuş katmanlarda, sınıflandırmada kullanılacak özellikler ağız öğrenmesi ile elde edilmektedir. Derin öğrenme algoritmaları; kendilerine sağlanan veriyi doğrudan öğrenmek yerine, hangi bilginin öğrenileceğine kendisi karar vermektedir. Bu sayede söz konusu algoritmalar klasik yöntemlere göre daha başarılı sonuçlar sunmaktadır (Kızrak ve Polat, 2018). Literatürde derin öğrenme için kullanılan 3 çeşit yapay sinir ağı türü bulunmaktadır. Bunlar; evrimsel sinir ağları (CNN), tekrarlayan sinir ağları (RNN), uzun-kısa vadeli hafıza ağları (LSTM)'dir. Çok katmanlı algılayıcıların (MLP) bir türü olan evrimsel sinir ağları, çoğunlukla görüntü işleme çalışmalarında kullanılmaktadır. CNN ağları evrim katmanları ve yardımcı ek katmanların bir araya gelmesi ile oluşmaktadır. Söz konusu katmanlardan evrim katmanı, evrimsel sinir ağlarının ana yapı taşıdır. Düğümler arasındaki bağlantıların yönlendirilmiş bir döngü oluşturduğu tekrarlayan sinir ağlarında (RNN) ise temel amaç; ardışık bilgileri kullanmaktır. Tekrarlayan sinir ağları çoğunlukla doğal dil işleme (NLP) işlemlerinde kullanılmaktadır. NLP işlemlerinde, bir cümle içindeki bir sonraki kelimeyi tahmin

edebilmek için daha önce hangi kelimelerin geldiğinin bilinmesi gerekmektedir. Tekrarlayan sinir ağları, cümledeki her bir kelime için önceki çıktıları değerlendirerek tahmin işlemi gerçekleştirmesinden dolayı NLP işlemlerinde yüksek başarı oranı sağlamaktadır (Kutlu, 2019). Tekrarlayan sinir ağları bazı durumlarda geçmişten gelen bilgiyi kullanmakta zorlanmaktadır. Bu durumlar için çözüm sağlayan uzun-kısa vadeli hafıza ağları (LSTM) Hochreiter ve Schmidhuber tarafından geliştirilmiştir (Hochreiter ve Schmidhuber, 1997). Uzun-kısa vadeli hafıza ağları, konuşma tanıma ve metin işleme gibi işlemlerde oldukça yüksek başarı oranı sağlamaktadır (Kutlu, 2019).

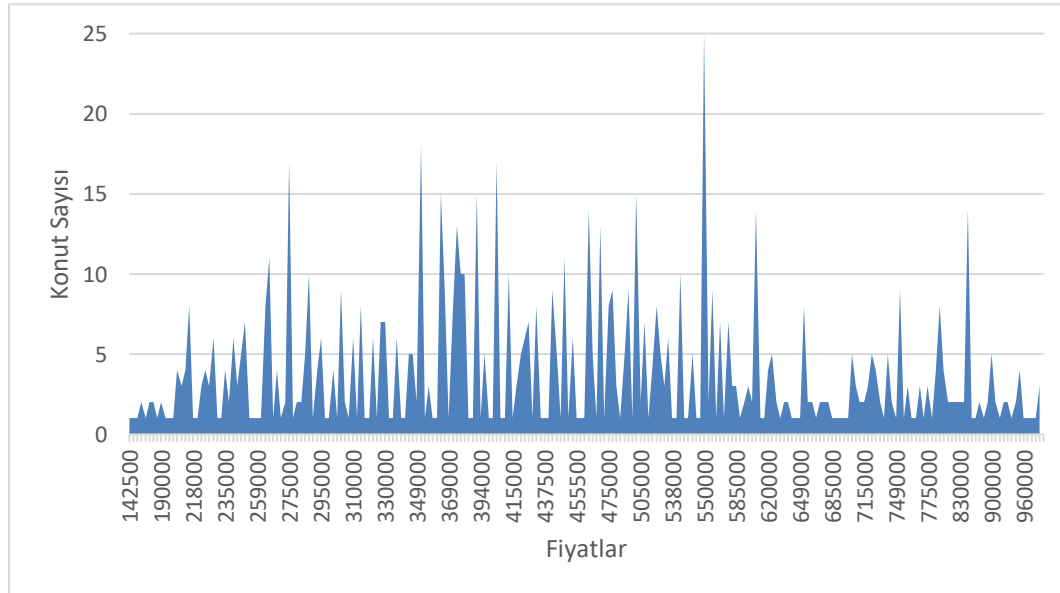
Bulgular

İstanbul ilinin Ataşehir ilçesine ait 852 adet konut verisi çalışma kapsamında internet ortamından toplanmıştır. Konut fiyatları 142.500 TL ile 1.000.000 TL arasında değişiklik göstermektedir. Toplanan verilere ait özellikler Tablo 1’de gösterilmiştir.

Tablo 1: Toplanan Veriler ve Özellikleri

Özellik	Değeri
Konut Sayısı	852
Ortalama Fiyat	472.196 TL
Fiyat Standart Sapma	187.702 TL
Minimum Fiyat	142.500 TL
Maksimum Fiyat	1.000.000 TL

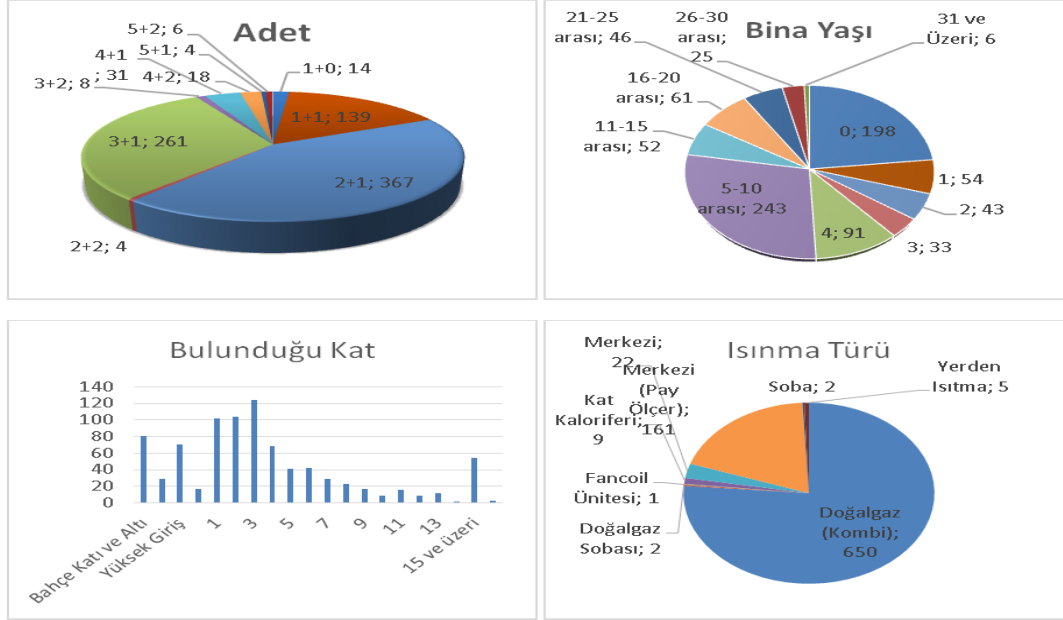
Konutlar brüt olarak 37 m2 ile 300 m2 arasında değişiklik göstermektedir. Toplam 16 mahalleye ait konut fiyatları 300.000 TL ile 600.000 TL arasında yoğunlaşırken bir milyon liradan daha fazla fiyata sahip konutlar kapsam dışı bırakılmıştır. Konut fiyatları ve fiyat sayıları Şekil 2’de gösterilmiştir.



Şekil 2: Konut Fiyatları ve Sayıları

Konutlar çoğunluğu 2+1 ve 3+1 dairelerden oluşmakta olup bina yaşları en çok sıfır ya da 5-10 yaş arasındadır. Satılık daireler arasında çoğunlukla 1-5 aralığındaki daireler ile bahçe katı ve

altındaki daireler vardır. Dairelerin büyük çoğunluğu doğalgaz kombi ile ısınırken kalanlar arasındaki büyük bir kısmı da ısı pay ölçer sistemine sahiptir. Konutların genel özellikleri aşağıdaki Şekil 3'te gösterilmiştir.



Şekil 3: Konutların Genel Özelliklerinin Dağılımı

Ataşehir ilçesindeki konut fiyatları Weka programı kullanılarak çeşitli algoritmalar ile tahmin edilmiştir. Yapılan analizler sonucu en başarılı sonuç Rastgele orman algoritması ile elde edilmiştir. Veriler 10-katlı çapraz doğrulama yöntemi ile eğitim ve test olacak şekilde ayrıştırılmıştır. Konut fiyatları ortalama ± 80 bin TL farklar ile tahmin edilebilmektedir. Gerçek fiyatlar ile tahmin edilen fiyatlar arasındaki korelasyon katsayısı 0,86 olarak bulunmuştur. Bu da ikisi arasında yüksek bir ilişki olduğunun göstergesi olup başarılı bir tahmin elde edildiğini göstermektedir. Aşağıdaki Tablo 2’de farklı algoritmaların kullanımı ile başarı sonuçları gösterilmiştir.

Tablo 2: Konut Fiyat Tahminlerinde Kullanılan Farklı Algoritmalar ve Sonuçları

Kullanılan Yöntem	Korelasyon Katsayısı	Ortalama Hata	Mutlak Hata	Ortalama Hata Karelerinin Karekökü
trees.DecisionStump	0,5838	119883,18		150302,10
trees.M5P	0,7952	72146,62		117892,43
trees.RandomForest	0,8608	80207,46		105105,38
trees.RandomTree	0,6126	119052,66		160776,30
trees.REPTree	0,7154	99436,63		132617,03
rules.DecisionTable	0,7488	90980,33		123655,51
rules.M5Rules	0,6931	73120,94		157577,15
rules.ZeroR	-0,0861	149503,08		185288,68
lazy.Ibk	0,605	118864,10		160854,94
lazy.KStar	0,7371	95781,80		130162,97
lazy.LWL	0,6174	115753,12		145850,35
functions.GaussianProcesses	0,8005	82001,74		114317,12
functions.LinearRegression	-0,0861	149503,08		185288,68
functions.SMOreg	0,7188	93853,53		144569,88

Sonuç ve Değerlendirme

Tarih boyunca konutların temel görevi insanları fiziksel etkenlerden korumak olmuştur. Günümüzde konutlar, insanların barınma ihtiyacını karşılamasının yanında, yatırım yapma ve kazanç sağlama imkânı da sağlamaktadır. Konutlar ayrıca insanların sosyal ve kültürel sınıflarını en net yansıtan mekânlar halini almışlardır. Söz konusu etmenler konut alım satım piyasasının ciddi oranda büyümesine ile sonuçlanmıştır. Bunun yanında aile birey sayısının artması, mevcut konutların ihtiyaçlara yetersiz kalması gibi etmenler insanların yeni konut arayışına girmesine neden olmuştur.

Günümüzde konut alım sürecinde konutların sahip olduğu özellikler kadar, fiyatı da büyük önem taşımaktadır. Konut fiyatları, serbest piyasa ekonomisi dolayısı ile değişkenlik göstermekte, aynı şehir içerisindeki benzer konutlar arasında dahi büyük fiyat farkları ortaya çıkabilmektedir. Bu nedenle konutlar için gerçekçi fiyatların belirlenmesi büyük önem taşımaktadır. Çünkü insanlar yıllar boyu yapmış oldukları birikimleri, ya da gelecekte elde edecekleri kazançların büyük kısmını alacakları konutlara yatırmaktadırlar. Bu, günümüzde birçok aile için bir kez yapılacak ve hata yapılma lüksü olmayan bir işlemdir. Bu denli önemli kararın alınmasında, alınacak konuta göre doğru bir fiyat tahmininde bulunan sistemler büyük önem arz etmektedir.

Bu çalışmada, İstanbul ilinin Ataşehir ilçesinde yer alan satılık daire ilanları toplanarak fiyat tahmin işlemi yapan bir sistem geliştirilmiştir. Gerçekleştirilen işlemde ilk olarak söz konusu ilçeye ait 852 adet satış ilanı internet üzerinden toplanmıştır. Elde edilen ilanlar analiz edilerek, metre kare, cephe, ısıtma, kat, cephe, oda sayısı, banyo sayısı, balkon sayısı, kullanım durumu gibi 176 adet farklı öznitelik içeren bir veri seti oluşturulmuştur. Gerçekleştirilen çalışmada kullanılan veri seti incelendiğinde; konut fiyatlarının 142.500 ile 1.000.000 TL arasında değiştiği görülmüştür. Her fiyat değerine sahip örneğin bulunması, farklı özelliklere sahip konutlar içinde fiyat tahmin işleminin gerçekleştirilebilmesine olanak sağlamıştır. Söz konusu çalışmada kullanılan veri seti hem elde edilme biçimi hem de sahip olduğu öznitelik çeşitliliği bakımından literatüre özgün bir değer kazandırmaktadır.

Gerçekleştirilen çalışmada oluşturulan veri seti kullanılarak, Weka programı aracılığı ile konut fiyat tahmin işlemleri gerçekleştirilmiştir. Söz konusu işlemde 14 farklı tahmin modeli kullanılmış, modellerden elde edilen başarı oranları karşılaştırılmıştır. Tahmin işlemlerinde, eğitim ve test veri setleri, 10 katlı çapraz doğrulama yöntemi kullanılarak elde oluşturulmuştur. Böylelikle tahmin işlemlerinden elde edilen sonuçların geçerliliği ve güvenilirliği arttırılmıştır. Farklı algoritmalar kullanılarak gerçekleştirilen fiyat tahmin işlemlerinde en yüksek başarı, toplu sınıflandırma algoritmalarından Rastgele Orman algoritmasından elde edilmiştir. Rastgele orman algoritması kullanılarak elde edilen tahmin verileri ile gerçek konut fiyatları karşılaştırıldığında, korelasyon katsayısı 0,86 olarak elde edilmiştir.

Literatürde konut fiyatları konusunda gerçekleştirilen çalışmaların birçoğu konutların sahip oldukları özelliklerin fiyata etkisi üzerine gerçekleştirilmiştir. Kördiş vd. (2014), Uyar ve Yayla (2016), Hayrulloğlu vd. (2017) ve Gökler (2017) gerçekleştirdikleri çalışmalarında konut özelliklerinin konut fiyatına olan etkilerini araştırmış, farklı bölge ve şehirler üzerinde yorumlamışlardır. Ekşioğlu ve Çubukçu (2011) tarafından gerçekleştirilen çalışmada ise, konut özelliklerinin satın alım ve kiralama seçenekleri üzerindeki etkisi incelenmiştir. Gerçekleştirilen bu çalışmada ise, söz konusu çalışmalardan farklı olarak konut özellikleri, konut fiyatlarının tahmin edilmesinde kullanılmıştır.

Literatürde konut özellikleri üzerinden fiyat ve maliyet tahmini için gerçekleştirilmiş çeşitli çalışmalar da bulunmaktadır. Erdem (1996) çalışmada, konutların inşa edilmesi öncesinde kullanılacak, inşa maliyetinin tahmin edilmesi için bir model geliştirmiştir. Demirel (2007) çalışmada ise, çok katlı toplu konut projelerinin inşaat maliyetlerinin hesaplanması için yapay sinir ağları yöntemi kullanarak bir model geliştirmiştir. Söz konusu çalışmalar, konutlar henüz proje aşamasında iken inşa için gerekli maliyetin hesaplanması ve inşa sürecinin planlanması amacı ile gerçekleştirilmiştir. Bu çalışma kapsamında geliştirilen model, inşa sonrasındaki satın alma

sürecinde gerçekçi fiyatların belirlenmesinde kullanılmaktadır. Uğur vd. (2011) çalışmalarında; konut maliyetlerini 21 farklı yığma konut projesine ait verileri kullanarak tahmin etmişlerdir. Söz konusu çalışmada hata oranı %5 olarak elde edilmiştir. Bu çalışma kapsamında geliştirilen model ise 825 örnek içeren bir veri seti kullanılarak geliştirilmiş ve hata oranı %14 olarak belirlenmiştir. Gerçekleştirmiş olduğumuz çalışmanın hata oranı her ne kadar diğer çalışmaya göre yüksek olsa da model geliştirilmesinde kullanılan örnek sayıları arasında ciddi bir fark görülmektedir. Bu çalışma kapsamında geliştirilen model, yüksek sayıda örnek içeren veri seti kullanılarak geliştirildiğinden dolayı, farklı türde test örnekleri için daha doğru fiyat tahmini sunmaktadır. Yılmaz vd. (2018) çalışmalarında, Eskişehir ilindeki satılık konut ilanlarından elde etmiş oldukları verileri kullanarak, konut fiyat tahmini gerçekleştiren yapay sinir ağları tabanlı bir model geliştirmişlerdir. Geliştirilen model kullanılarak gerçekleştirilen tahmin işlemlerinde korelasyon değeri 0.9219 olarak tespit edilmiştir. Söz konusu çalışmada kullanılan veri setinde toplam 19 öznelik bulunmaktadır. Bu çalışma kapsamında kullanılan veri seti içerisinde ise konut özelliklerini ifade eden 176 öznelik bulunmaktadır. Günümüzde konutlar fiyatlarına etki eden birçok farklı özelliğe sahiptir. Söz konusu özellikler alım satım sürecinde insanların kararı üzerinde önemli rol oynayabilmektedir. Bu sebeple olabildiğince çok özelliği parametre olarak alan modeller, insanların talepleri doğrultusunda daha doğru fiyat tahmininde bulunacaktır.

Elde edilen sonuçlar incelendiğinde, geliştirilen model kullanılarak konutlar için başarılı fiyat tahminlerinin yapılabildiği görülmüştür. Geliştirilen tahmin sisteminin, konut alırken fiyat belirlemede sorun yaşayan insanlara, doğru fiyat önerisi sunmak amacı ile kullanılabileceği düşünülmektedir.

Kaynakça

- Afşar, A., Yılmazel, Ö., & Yılmazel, S. (2017). Konut Fiyatlarını Etkileyen Faktörlerin Hedonik Model İle Belirlenmesi: Eskişehir Örneği. *Selçuk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, (37), 195-205.
- Akar, Ö., Güngör, O. (2013). Eş Dizimlilik Matrisi ve Rastgele Orman Sınıflandırıcısı ile Çay ve Fındık Alanlarını Sınıflandırılması. *Türkiye Ulusal Fotogrametri ve Uzaktan Algılama Birliği VII. Teknik Sempozyumu*. 23-25 Mayıs, Trabzon, Türkiye.
- Akar, Ö., Güngör, O., Akar, A. (2010). Rastgele Orman Sınıflandırıcısı ile Arazi Kullanım Alanlarının Belirlenmesi, *III. Uzaktan Algılama ve Coğrafi Bilgi Sistemleri Sempozyumu*, 11-13 Ekim 2010, Gebze, Kocaeli, Bildiriler Kitabı, pp.142-152.
- Akar, Ö., ve Güngör, O. (2012). Rastgele Orman Algoritması Kullanılarak Çok Bantlı Görüntülerin Sınıflandırılması. *Jeodezi ve Jeoinformasyon Dergisi*, 1(2), 139-146.
- Aktaş, A. (2020). *Derin Öğrenme Yöntemleri İle Dokunsal Parke Yüzeyi Tespiti*, Marmara Üniversitesi, Fen bilimleri Enstitüsü. Yüksek Lisans Tezi, İstanbul, 61s.
- Arslan, Y., Ceritoğlu, E., & Kanık, B. (2013). The Effects Of Demographic Changes On The Long Term Housing Demand in Turkey. *Munich Personal Repec Archive*.
- Bostancıoğlu, E. (1999). Konut binalarının ön tasarımı evresinde maliyeti etkileyen faktörler ve faktörlere dayalı bir maliyet tahmin yöntemi (Doctoral dissertation, Fen Bilimleri Enstitüsü).
- Bostancıoğlu, E. (2006). Konut Binalarının Ön Tasarımında Bir Maliyet Tahmin Modeli. *Dokuz Eylül Üniversitesi Mühendislik Fakültesi Fen ve Mühendislik Dergisi*, 8(3), 27-49.
- Demirel, Y. (2007). Toplu Konut İnşaat Maliyetlerinin Yapay Sinir Ağları İle Tahmini. *Selçuk Üniversitesi Mühendislik, Bilim ve Teknoloji Dergisi*, 22(4), 53-60.

- Ecer, F. (2014). Türkiye'deki Konut Fiyatlarının Tahmininde Hedonik Regresyon Yöntemi İle Yapay Sinir Ağlarının Karşılaştırılması. *In International Conference On Eurasian Economies*, (pp. 1-10).
- Ekşioğlu Çetintahra, G., & Çubukçu, E. (2011). Çevre Estetiğinin Konut Fiyatlarına Etkisi. *İtû Dergisi/A*, 10(1), 3-12.
- Ercire, M. (2019). *Kısa süreli güç kalitesi bozulmalarının dalgacık analizi ve rastgele orman yöntemi ile sınıflandırılması* (Master's thesis, Kütahya Dumlupınar Üniversitesi/Fen Bilimleri Enstitüsü), Yüksek Lisans Tezi, Kütahya, 111s.
- Erdem, A. E. (1996). *Konut Projelerinde, Ön Tasarım Aşamasında, Fonksiyonel Elemanlara Dayalı Maliyet Tahmini İçin Bir Model Geliştirilmesi*, İstanbul Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi.
- Fitöz, E., & Öztürk, N. (2008). *Türkiye'de Konut Piyasasının Belirleyicileri: Ampirik Bir Uygulama*. Yayınlanmamış Yüksek Lisans Tezi, Zonguldak Karaelmas Üniversitesi, SBE.
- Gökler, L. A. (2017). Ankara'da Konut Fiyatları Farklılaşmasının Hedonik Analiz Yardımıyla İncelenmesi. *Megaron*, 12(2), 304-315.
- Hayrullahoğlu, G., Aliefendioğlu, Y., Tanrıvermiş, H., & Hayrullahoğlu, A. C. (2017, September). Konut Piyasalarında Hedonik Değerleme Modeli Tahmini: Ankara İli Çankaya İlçesi Çukurambar Bölgesi Örneği. *In Proceedings of 2nd International Conference on Scientific Cooperation for The Future in The Economics and Administrative Sciences*, (P. 25).
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8), 1735-1780.
- Kızrak, M. A., ve Bolat, B. (2018). Derin öğrenme ile kalabalık analizi üzerine detaylı bir araştırma. *Bilişim Teknolojileri Dergisi*, 11(3), 263-286.
- Korkem, E. (2013). *Mikroarray Gen Ekspresyon Veri Setlerinde Random Forest Ve Naive Bayes Sınıflama Yöntemleri Yaklaşımı*, Hacettepe Üniversitesi, Sağlık Bilimleri Enstitüsü, Yüksek Lisans Tezi.
- Kördiş, G., Işık, S. & Mert, M. (2014). Antalya'da Konut Fiyatlarını Etkileyen Faktörlerin Hedonik Fiyat Modeli İle Tahmin Edilmesi. *Akdeniz Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 14(28), 103-132.
- Kutlu, Ö. (2019). *İnsansız Hava Aracı İle Elde Edilen Görüntülerin Derin Öğrenme Yöntemleri İle Analizi*, Marmara Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi.
- Lebe, F., & Akbaş, Y. E. (2014). Türkiye'nin Konut Talebinin Analizi: 1970-2011. *Ataturk University Journal Of Economics & Administrative Sciences*, 28(1), 57-83.
- LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature*, 521, 436-444.
- Marın, M. C., & Altıntaş, H. (2004). Konut Yer Seçimi-Ulaşım Etkileşim Teorileri: Kritik Bir Literatür İncelenmesi. *Gazi Üniversitesi Mühendislik-Mimarlık Fakültesi Dergisi*, 19(1), 73-88.
- Türel, A. (1976). Ankara'da Konut Fiyatlarının Mekânsal Farklılaşması. *Geoforum*, 7, 271-283.
- Uğur, L. O., & Sivri, A. R. (2014). Toplu Konut Projelerinde Maliyetlerin Regresyon Metodu İle İncelenmesi. *Düzce Üniversitesi Bilim ve Teknoloji Dergisi*, 2(1), 251-270.
- Uğur, L. O., Baykan, U. N., & Korkmaz, S. (2011). Yığma Konutların Maliyet Tahmininde Yapay Sinir Ağlarının (Ysa) Kullanılması, 6. *İnşaat Yönetim Kongresi*, Bursa.

-
- Uyar, S. G. K., & Yayla, N. (2016). Konut Fiyatlarının Hedonik Fiyatlama Yaklaşımına Göre Mekânsal Ekonometrik Modeller İle Tahmini: İstanbul Konut Piyasası Örneği. *Social Sciences*, 11(4), 326-342.
- Yayar, R., & Gül, D. (2014). Mersin Kent Merkezinde Konut Piyasası Fiyatlarının Hedonik Tahmini. *Anadolu University Journal Of Social Sciences*, 14(3), 87-99.
- Yayar, R., & Karaca, S. S. (2014). Konut Fiyatlarına Etki Eden Faktörlerin Hedonik Modelle Belirlenmesi: Tr83 Bölgesi Örneği. *Ege Academic Review*, 14(4), 509-518.
- Yılmazel, Ö., Afşar, A., & Yılmazel, S. (2018). Konut Fiyat Tahmininde Yapay Sinir Ağları Yönteminin Kullanılması. *Uluslararası İktisadi ve İdari İncelemeler Dergisi*, 20, 285-300.